

Сравнительный анализ алгоритмов построения больших коммуникационных сетей со свойствами “малого мира”

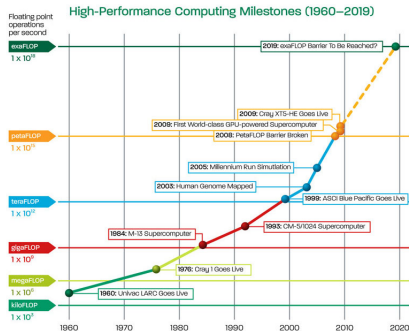
А. Демичев, В. Ильин, А. Крюков, С.Поляков

НИЦ “Курчатовский институт” и НИИЯФ МГУ

Суперкомпьютерный форум
“СУПЕРКОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ В
ОБРАЗОВАНИИ, НАУКЕ И ПРОМЫШЛЕННОСТИ”
НРС'2012

Суперкомпьютеры следующего поколения

- ▶ Появление реальных вычислительных систем следующего поколения - с производительностью порядка эксафлопса (10^{18} FLOPS) - ожидается не ранее 2018–2020 года
- ▶ однако принципы их построения начинают интенсивно разрабатываться уже сейчас
- ▶ на пути к их построению предстоит решить ряд сложных научно-технических задач
- ▶ выработать принципиально новые решения для их архитектуры и аппаратной реализации.



Суперкомпьютеры следующего поколения: коммуникационная сеть

- ▶ Коммуникационная сеть - одна из важнейших составляющих любого суперкомпьютера
 - ▶ в первую очередь определяет возможность увеличения числа вычислительных узлов, что необходимо для достижения желаемой производительности
- ▶ Как ожидается, коммуникационная сеть суперкомпьютеров экзафлопного уровня будет объединять $> 100\,000$ вычислительных узлов (“end points”)
- ▶ \Rightarrow разработка коммуникационных сетей с хорошими свойствами масштабируемости и возможностью эффективно обслуживать огромное число вычислительных узлов



Задача проекта

- ▶ разработать подход к построению и оптимизации коммуникационных сетей общего вида — с неоднородной структурой
- ▶ в настоящее время проект находится в начальной стадии:
 - ▶ формулировка принципов подхода
 - ▶ выбор методов построения и исследования архитектуры сетей
 - ▶ разработка инструментария для численного моделирования
 - ▶ предварительные первые результаты
- ▶ частично финансируется РФФИ, грант 12-07-00408-а.

Исходные принципы подхода

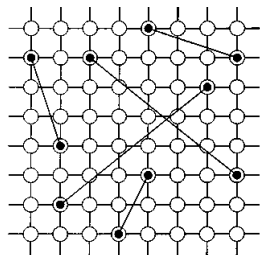
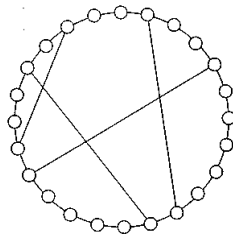
- ▶ Сеть должна обладать свойствами “малого мира”, а именно, медленным (логарифмическим) ростом среднего расстояния между узлами с ростом их числа
 - ▶ \implies масштабируемость
- ▶ Сеть должна иметь базовую структуру D -мерной регулярной решетки
 - ▶ Например, для моделирования D -мерных объектов
- ▶ Эта базовая структура должна иметь дополнительные связи таким образом, чтобы могли достаточно оптимальным образом решаться и общие задачи \rightarrow малая средняя сетевая длина коммуникаций, оптимальная навигация и маршрутизация по сети
 - ▶ Желательно, чтобы характеристики новых связей зависели от варьируемых параметров для возможности оптимизации

Решеточные сети с перемычками (PCП) - модификация алгоритма Уоттса-Строгача (Watts-Strogatz, 1998)

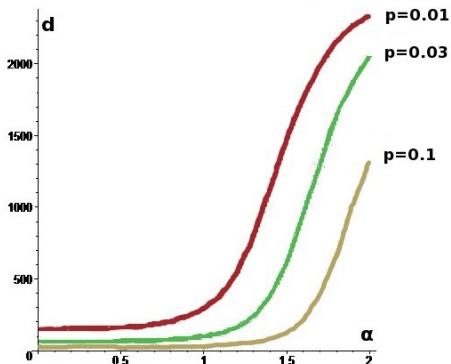
1. исходный объект - D -мерная регулярная решетка с топологией тора, содержащая $N = L^D$ узлов ;
2. последовательно перебираются все узлы решетки и с вероятностью $0 < p \leq 1$ к каждому узлу i подсоединяют первый конец перемычки;
3. второй конец перемычки выбирается случайно с вероятностью

$$P(r) \sim r^{-\alpha}$$

- функция решеточного расстояния r между узлами



Зависимость средней длины от параметра распределения
длин перемычек $P(r) \sim r^{-\alpha}$ и вероятности перемычек p
 $L = 10^4$ $D = 1$



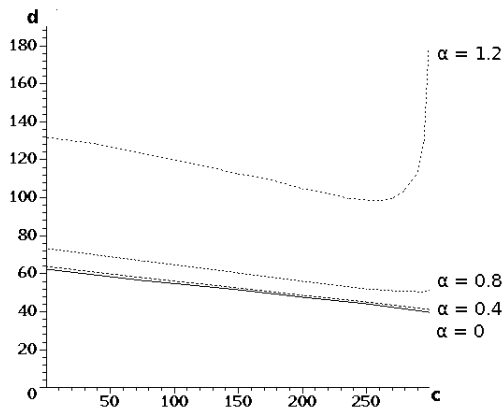
- ▶ Увеличение среднего пути между узлами с ростом параметра α
- ▶ Наивный вывод: надо использовать сети с минимальным α

Уменьшение степени стохастичности РСП

- ▶ Коммуникационным сетям СК внутренне не присуща стохастичность \Rightarrow может быть удобно использовать модификацию базового алгоритма:
 - ▶ фиксируется не вероятность p появления перемычки из каждого узла, а полное число перемычек t , которые должны быть добавлены к решетке;
- ▶ Дальнейшее уменьшение степени стохастичности РСП
- ▶ $c \leq t$ перемычек добавляются специальным образом: их концы соединены
- ▶ В определенной степени этот алгоритм является гибридным: частично стохастическим, а частично детерминистским

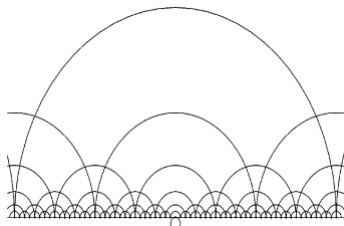
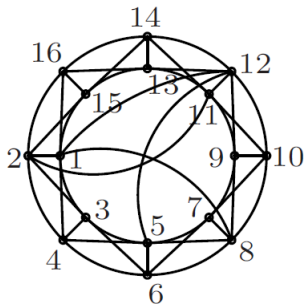
Уменьшение степени стохастичности РСП (2)

- ▶ Среднее расстояние уменьшается с ростом числа заведомо соединенных между собой перемычек
- ▶ \Rightarrow с точки зрения средней длины этот алгоритм является предпочтительным по сравнению с исходным стохастическим алгоритмом



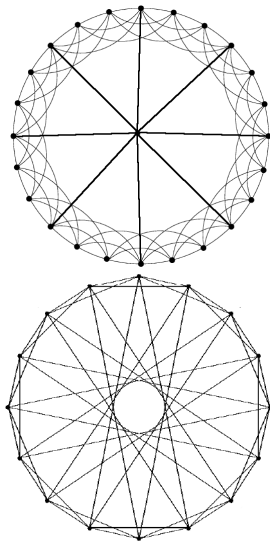
РСП без стохастичности: детерминистские алгоритмы

- ▶ Самый простой и естественный алгоритм: на каждом шаге алгоритма осуществляется соединение перемычкой наиболее удаленных друг от друга в сетевом смысле узлов
- ▶ Иерархический алгоритм HN4 основан на представлении номеров узлов $0 < n < L$:
$$n = 2^i(2j + 1)$$



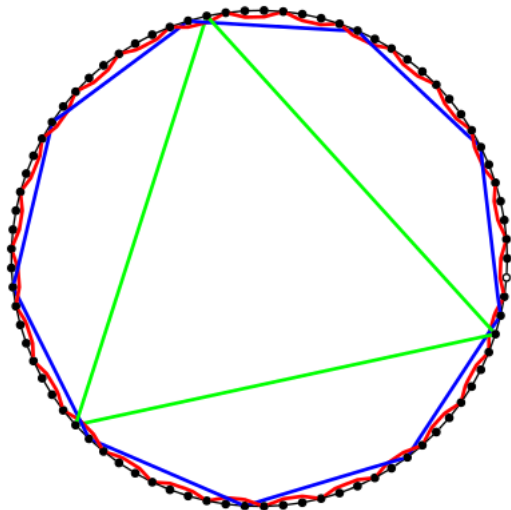
РСП без стохастичности: детерминистские алгоритмы (2)

- ▶ Выделение узлов-хабов, которые соединяются ребрами (перемычками) так, что получается некоторый граф H со специальными свойствами
- ▶ Мультипликативные циркулянтные графы $C(s^k; 1, s, s^2, \dots, s^{k-1})$



РСП без стохастичности: субциркулянтные (хордовые) сети

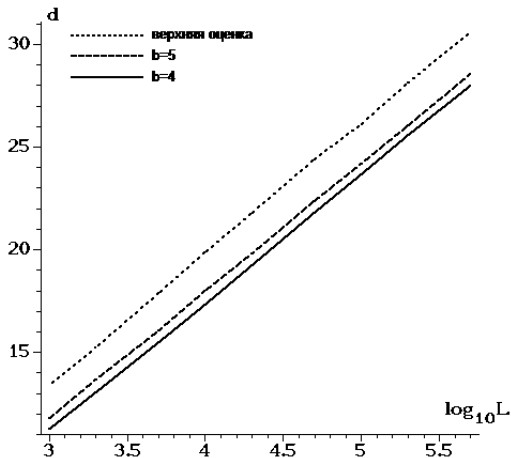
- ▶ узлы $0, b, 2b, \dots, L - b$ соединяются перемычками в цикл (узел 0 соединяется с b , b с $2b$ и т.д., $L - b$ с 0);
- ▶ для каждого $i = 2, \dots, k$ узлы $0, b^i, 2b^i, \dots, L - b^i$ соединяются перемычками в цикл



Субциркулянтные сети \leftrightarrow малый мир

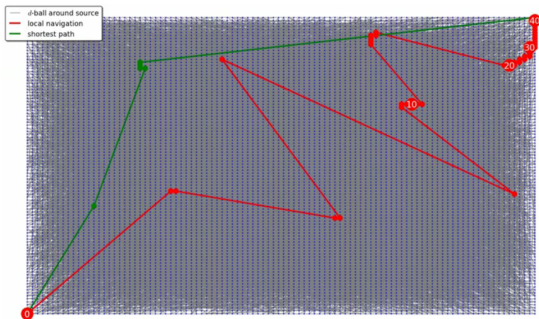
- ▶ приближенная верхняя оценка:

$$d \lesssim \lambda \left[\frac{k(b+4)}{2} + \frac{L}{4b^k} - 2 \right] \quad \text{где } \lambda \simeq 1$$



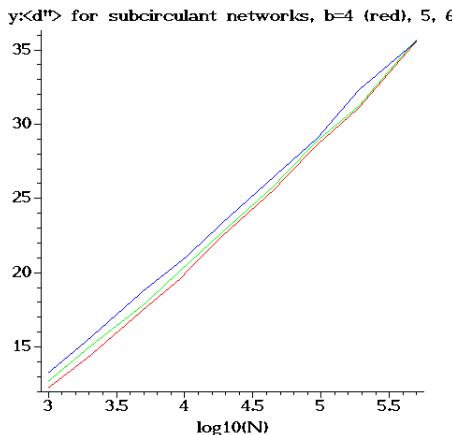
Навигация при наличии только локальной информации

- ▶ сообщение передается от узла к узлу по ребрам (связям);
- ▶ узел "знает" географическое положение (другими словами, положение в основной решетке) всех узлов и своих ближайших сетевых соседей;
- ▶ необходимо доставить сообщение в узел назначения по возможно кратчайшему пути.
 - ▶ *жадный алгоритм*: данный узел пересылает сообщение тому своему соседу, который географически ближе всего к цели (узлу назначения).



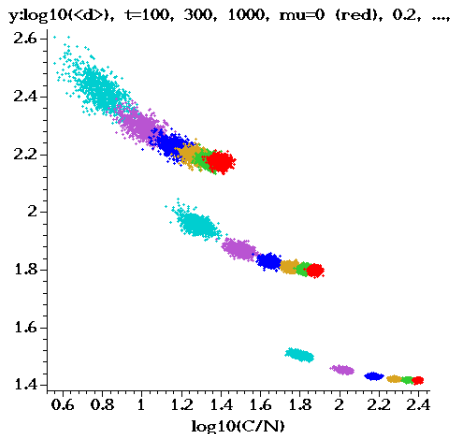
Навигационная длина пути в иерархических РСП

- ▶ Иерархические РСП
 - ▶ весьма малое среднее расстояние d
 - ▶ большая навигационная средняя длина пути ℓ
- ▶ *двухуровневая* локальная навигация:
просматриваются не только ближайшие соседи, но и соседи соседей
 - ▶ субциркулянтная РСП: $\ell^{(2)}$ от $\log L$ с $b = 4, 5, 6$ и k , минимизирующими $\ell^{(2)}$ для выбранных L



Учет “цены” построения РСП

- ▶ Необходимо учитывать “цену” построения сети; иначе - тривиальное решение = полный граф
- ▶ Естественная функция “цены” - удельная длина перемычек = (суммарная длина)/(число узлов)
- ▶ большое количество не слишком длинных перемычек является более дешевым решением.



Многокритериальная оптимизация “цена – качество”

- ▶ Оптимизации с векторной целевой функцией
- ▶ В нашем случае — два показателя:
 - ▶ “качество” РСП = среднее расстояние d или ℓ или $\ell^{(2)}$
 - ▶ “цена” построения РСП = удельная длина перемычек C/L
- ▶ метод взвешенных сумм \Rightarrow скалярные целевые функции

$$G_w = wd + (1 - w)C/L, \quad (1)$$

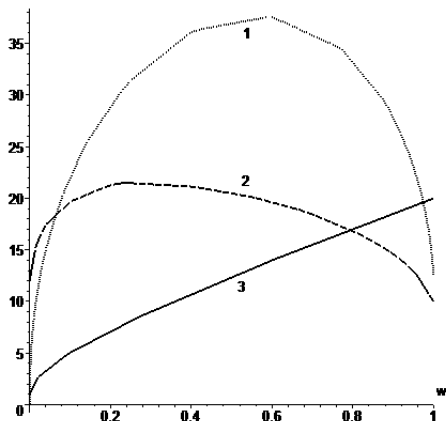
$$G'_w = w\ell + (1 - w)C/L, \quad (2)$$

$$G''_w = w\ell^{(2)} + (1 - w)C/L. \quad (3)$$

- ▶ параметр $0 \leq w \leq 1$ характеризует относительную значимость каждого из критериев (качество и цена)
- ▶ оптимизация = $\forall w$ подобраны значения параметров алгоритмов ($p/t, \alpha, c, b, k, \dots$): $\min G_w, G'_w$ или G''_w
 - ▶ размер сети L считаем заданным

Сравнение детерминистских и стохастических алгоритмов

- ▶ 1 — $\min G_w''$ для стохастических сетей с соединенными перемычками
- ▶ 2 — $\min G_w$ для детерминистских сетей с двухпетлевым (double-loop) соединением хабов
- ▶ 3 — $\min G_w''$ для субциркулянтных сетей
- ▶ \Rightarrow предпочтительными являются субциркулянтные сети



Заключение

- ▶ Предложено использовать сеть с нерегулярной структурой
 - ▶ малая средняя длина пути между вычислительными узлами достигается минимальным добавлением новых связей к регулярной решетке
- ▶ Рассмотрен ряд как известных в литературе, так и оригинальных алгоритмов построения сложных сетей со свойствами “малого мира”
- ▶ Предложена методика сравнения эффективности алгоритмов различных типов на основе оптимизации соотношения “цены” и “качества”
- ▶ Двумя основными классами алгоритмов являются стохастические и детерминистские алгоритмы ⇒ сравнение на основе соотношения “цена – качество”
- ▶ Предпочтительным типом сетей в широком диапазоне параметра относительной значимости указанных критериев являются субциркулянтные сети, которые строятся с помощью детерминистского алгоритма

Дальнейшие планы

- ▶ Рассмотрение свойств предложенных в данной работе сетей, как с точки зрения других характеристик (в частности, нагрузки на узлы, устойчивости и т.п.)
- ▶ Обобщения на более высокие размерности базовой решетки.
- ▶ Построение модели маршрутизации потоков сообщений; разработка
 - ▶ \Rightarrow аналитических моделей
 - ▶ \Rightarrow программного инструментария для численного моделирования потоков в решеточных сетях с перемычками и свойствами малого мира